

Amir - Amir
Aliq

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:35:16 ; Search time 30.13 Seconds

(without alignments)
409,200 Million cell updates/sec

Title: US-10-048-197-2

Perfect score: 558

Sequence: 1 MLNHHIRLTISALTALVT.....FLWPKYRPNLSFYLTAKA 111

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	558	100.0	111	22	AAB60640
2	78	14.0	300	21	AAW81582
3	74	13.3	112	21	AAH21107
4	74	13.3	112	21	AAV93269
5	69.5	12.5	334	21	AAH31852
6	69.5	12.5	335	21	AAH31851
7	69.5	12.5	337	21	AAH31850
8	68	12.2	322	21	AAH31124
9	68	12.2	323	21	AAH31123
10	68	12.2	325	21	AAH31122
11	67.5	12.1	151	17	AAH95910

12	66.5	11.9	222	22	ABG02873	Novel human diageno
13	66	11.8	421	22	AAH39816	Human polypeptide
14	66	11.8	475	22	AAH41602	Human polypeptide
15	64.5	11.6	281	22	AAE04356	Oryza sativa rdt1f
16	64.5	11.6	564	22	AAE04359	Glycine max aspart
17	64.5	11.6	2785	21	AAV57148	Human down-regulat
18	64	11.5	970	22	AAH67548	Amino acid sequenc
19	63.5	11.4	295	22	AAH92147	C. glutamicum prote
20	63.5	11.4	296	22	AAH92147	Corynebacterium gl
21	63.5	11.4	346	11	AAH06323	Human DNase. Homo
22	63.5	11.4	389	19	AAH96234	H. pylori GHP0 538
23	63.5	11.4	429	21	AAH54261	Arabidopsis thalia
24	63.5	11.4	430	21	AAH54260	Arabidopsis thalia
25	63.5	11.4	453	21	AAH54259	Arabidopsis thalia
26	63	11.3	196	22	AAH63940	Human prostate can
27	63	11.3	254	14	AAH33076	HpaI restriction e
28	63	11.3	981	20	AAH34476	Porphyromonas ging
29	63	11.3	985	20	AAH34352	Porphyromonas ging
30	63	11.3	1222	21	AAH36453	Arabidopsis thalia
31	63	11.3	1257	21	AAH36452	Arabidopsis thalia
32	63	11.3	1275	21	AAH36451	Arabidopsis thalia
33	62.5	11.2	255	22	AAH91022	C. glutamicum prote
34	62.5	11.2	363	22	AAH01104	Novel human diageno
35	62.5	11.2	782	22	AAH46720	ACNV DNA polymeras
36	62.5	11.2	999	22	AAH01105	Novel human diageno
37	62	11.1	326	21	AAH35229	Arabidopsis thalia
38	62	11.1	330	20	AAH35229	Chlamydia pneumoni
39	62	11.1	330	21	AAH29455	Arabidopsis thalia
40	61.5	11.0	214	22	AAH71901	Drosophila melanog
41	61.5	11.0	465	22	AAH66746	Drosophila melanog
42	61.5	11.0	787	22	AAH58186	Drosophila melanog
43	61.5	11.0	787	22	AAH66743	Drosophila melanog
44	61.5	11.0	787	22	AAH66744	Drosophila melanog
45	61.5	11.0	787	22	AAH66745	Drosophila melanog

ALIGNMENTS

RESULT 1	AAH60640	standard; Protein: 111 AA.
ID	AAH60640	
AC	AAH60640	
DE	02-MAY-2001	(first entry)
XX	Moraxella catarrhalis strain ATCC43617 BASB122 protein.	
XX	BASB122 protein: strain ATCC43617; antigen: antibody; vaccine;	
KW	genetic immunisation: infection: upper respiratory tract; otitis media;	
KW	hearing loss; deafness; pneumonia; sinusitis; nosocomial infection;	
KW	invasive disease; antibacterial; auditory.	
XX		
OS	Moraxella catarrhalis.	
PN	WO200109337-A2.	
PD	08-FEB-2001.	
PF	31-JUL-2000; 2000WO-EP07365.	
PR	30-JUL-1999; 99GB-0018034.	
PR	30-JUL-1999; 99GB-0018036.	
PI	(SMK) SMITHKLINE BEECHAM BIOLOGICALS.	
PI	Thomson J.	
DR	WPI: 2001-159874/16.	
XX	N-PSDB: AAH59778.	
PT	New BASB122 and BASB124 polypeptides and polynucleotides from Moraxella	

PT catarrhalis strain ATCC 43617, useful as therapeutic agents or vaccines
 PT against bacterial infections, e.g. otitis media or pneumonia
 XX
 PS Claim 4; Page 66; 75pp; English.
 CC The invention relates to the Moraxella catarrhalis strain ATCC43617
 CC BASB122 and BASB124 proteins (AAB60640 and AAB60641, respectively)
 CC and to DNA encoding them (AAFS9778 and AAFS9779, respectively). The
 CC invention also relates to immunogenic fragments of the BASB122 and
 CC BASB124 proteins, expression vectors and host cells comprising BASB122
 CC or BASB124 nucleic acids, the recombinant production of BASB122 or
 CC BASB124, vaccine compositions comprising the BASB122 or BASB124 proteins
 CC or nucleic acids, an antibody against BASB122 or BASB124, therapeutic
 CC compositions comprising an anti-BASB122 or anti-BASB124 antibody, and a
 CC method of identifying a Moraxella catarrhalis infection via the
 CC detection of BASB122 or BASB124 proteins or antibodies. The vaccine
 CC compositions of the invention are useful as prophylactic or therapeutic
 CC agents against Moraxella catarrhalis infections in mammals, particularly
 CC humans. Moraxella catarrhalis is a Gram negative bacterium frequently
 CC isolated from the human upper respiratory tract, which is responsible for
 CC several pathological conditions. It is responsible for about 15% of
 CC otitis media cases in children (which can lead to temporary or permanent
 CC hearing loss). It also causes pneumonia in elderly people, and sinusitis,
 CC nosocomial infections and, less frequently, invasive diseases. BASB122 or
 CC BASB124 proteins or nucleotides may additionally be used in screening for
 CC novel antibacterial compounds, and in the diagnosis and staging of
 CC infections. The present sequence represents the Moraxella catarrhalis
 CC strain ATCC43617 BASB122 protein.
 XX
 SQ Sequence 111 AA;
 Query Match 100.0%; Score 558; DB 22; Length 111;
 Best Local Similarity 100.0%; Pred. No. 3.9e-61;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLNHHIRLTISALTALIVGCVSTGNVAMKEONQOTIEQITIKGKTAKOISRSFGSAD 60
 Db 1 mlnhhirltisaltalivgcvstgnvamkeqngtiegikgktkgsifsgsad 60
 QY 61 SISPMIVIRKFGHTAIIAPNMOELSLISFLWVKPYRPNLSFYLAKA 111
 Db 61 sistmivirkfghtaiaipnmoelstlisflwvkpyrpnlsfytaka 111
 RESULT 2
 AAY81582
 ID AAY81582 standard; Protein: 300 AA.
 XX
 AC AAY81582;
 XX
 DT 24-MAY-2000 (first entry)
 XX
 DE Streptococcus pneumoniae type 4 protein sequence #82.
 XX
 KW Streptococcus pneumoniae; vaccine; screening; protein antigen;
 KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;
 KW pneumococcal disease.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200006737-A2.
 PD 10-FEB-2000.
 XX
 PE 27-JUL-1999; 99WO-GB02451.
 XX
 PF 27-JUL-1998; 98GB-0016337.
 PR 19-MAR-1999; 99US-0125164.
 XX
 PA (MICR-) MICROBIAL TECHNICS LTD.
 XX
 PI Gilbert CRG, Hansbro PM;

XX
 DR WPI; 2000-195300/17.
 XX
 PT New Streptococcal protein, useful as a vaccine, for diagnosis of
 PT pneumococcal diseases and for screening agents capable of antagonizing
 PT or inhibiting expression of the protein
 XX
 PS Claim 1; Page 80; 108pp; English.
 CC AAY81501 to AAY81679 represent specifically claimed protein sequences
 CC isolated from Streptococcus pneumoniae. AA05407 to AA05590 represent
 CC specifically claimed nucleotide sequences isolated from S. pneumoniae.
 CC The sequences have antibacterial and antiinflammatory properties.
 CC The protein sequences, and fragments of them, are useful as immunogens
 CC and/or antigens. The nucleotide sequences can be used in vaccines and in
 CC diagnostic assays. The proteins and nucleotides can be useful for the
 CC detection and diagnosis of S. pneumoniae. The protein sequences are also
 CC useful for screening an agent capable of antagonising, inhibiting or
 CC interfering with the function or expression of the proteins in which the
 CC agent is useful for treatment or prophylaxis of S. pneumoniae infection
 CC and meningitis. AA05591 to AA05614 represent primers used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 300 AA;
 Query Match 14.0%; Score 78; DB 21; Length 300;
 Best Local Similarity 28.3%; Pred. No. 0.36;
 Matches 36; Conservative 18; Mismatches 39; Indels 34; Gaps 6;
 QY 10 ISALLTA-LIVTGCVST--GNVAMKEONQOTIEQITII-----KGTN 48
 Db 78 lgalvtavilvtsvlienvtkllhpqvndegilwlgialtlnllaalvvgkgytk 137
 QY 49 KOEISSRFGSADSISEF-----MIWIKFGHTAIIAPNMOELSLISFLWVKPYRPNL 103
 Db 138 neslshlfedltgwwavilmavilrftdwyildp-----lslvstffllskalpr-- 190
 QY 104 SFYLTAK 110
 Db 191 -fwstlk 196
 RESULT 3
 AAB21107
 ID AAB21107 standard; Protein: 112 AA.
 XX
 AC AAB21107;
 XX
 DT 19-JAN-2001 (first entry)
 XX
 DE Neisseria meningitidis BASB059 protein.
 XX
 KW BASB059: respiratory tract infection; invasive bacterial disease;
 KW bacteraemia; meningitis.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO2000044904-A1.
 PD 03-AUG-2000.
 XX
 PE 25-JAN-2000; 2000WO-EP00561.
 XX
 PR 29-JAN-1999; 99GB-0002070.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Thonnard J;
 XX
 DR WPI; 2000-505978/45.
 DR N-PSDB; AAA76029.

PT New isolated polypeptide from *Neisseria meningitidis* is useful for
 detection and treatment of *N. meningitidis* infection -

PS Claim 3; Page 61; 77pp; English.

XX The present sequence is the *Neisseria meningitidis* BASB059 protein. The
 CC protein, its gene, agonists and antagonists can be used in diagnostic
 CC assays for the bacterium, as vaccines to prevent infection and as
 CC treatments for bacterial infection, particularly those caused by
 CC *Neisseria meningitidis*, such as upper respiratory tract infections,
 CC invasive bacterial diseases, bacteraemia and meningitis.

XX Sequence 112 AA:

Query Match 13.3%; Score 74; DB 21; Length 112;

Best Local Similarity 31.2%; Pred. No. 0.3;

Matches 20; Conservative 10; Mismatches 34; Indels 0; Gaps 0;

OY 7 RLTIALTLALTAVTGCSTGNVAMKEONQOTIEQTIKGTNKKEISRFSGADSISPMI 66
 DB 4 RLVSAFAVALAGCGSINNVTSDQKIGERAFAIYVGNVAKISNINSIINFIA 63

OY 67 VVIR 70
 DB 64 tvrk 67

RESULT 4

AA93269
 ID AAY93269 standard; Protein: 112 AA.

XX AAY93269;

DT 04-SEP-2000 (first entry)

XX Amino acid sequence of a polypeptide of a *Neisseria* pathogenic strain.

XX Pathogenic strain; *Neisseria*; vaccine; *Neisseria* infection.

XX *Neisseria meningitidis*.

XX WO200026375-A2.

XX 11-MAY-2000.

XX 28-OCT-1999; 99WO-FR02643.

XX 30-OCT-1998; 98FR-0013693.

PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
 (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PI Anjame L, Bouchardon A, Renauld-Mongenie G, Rokhl B, Nassif X;

PI Tinsley C., Perrin A;

DR WPI: 2000-365622/31.

DR N-PSDB; AAA15300.

PT New polypeptide specific for pathogenic *Neisseria* useful in therapeutic
 or preventative vaccines and for diagnosis -

PS Claim 5; Page 69-70; 187pp; French.

XX The present sequence represents a protein that is specific for pathogenic
 CC strains of *Neisseria*. The polynucleotides, polypeptides, or their
 CC antigenic fragments, are used in vaccines to treat or protect against
 CC *Neisseria* infections, particularly by *N. meningitidis*. The
 CC polynucleotide sequence is also used for recombinant production of
 CC the polypeptide and to produce attenuated *Neisseria* strains that
 CC overexpress it, or express it in a non-toxic mutant form.

XX Sequence 112 AA:

Query Match 13.3%; Score 74; DB 21; Length 112;
 Best Local Similarity 31.2%; Pred. No. 0.3;
 Matches 20; Conservative 10; Mismatches 34; Indels 0; Gaps 0;

OY 7 RLTIALTLALTAVTGCSTGNVAMKEONQOTIEQTIKGTNKKEISRFSGADSISPMI 66
 DB 4 RLVSAFAVALAGCGSINNVTSDQKIGERAFAIYVGNVAKISNINSIINFIA 63

OY 67 VVIR 70
 DB 64 tvrk 67

RESULT 5

AA631852
 ID AAG31852 standard; Protein: 334 AA.

XX AAG31852;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 38321.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 06-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132486.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 14-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134768.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

XX 24-MAY-1999; 99US-0135629.

XX 25-MAY-1999; 99US-0136021.

XX 27-MAY-1999; 99US-0136392.

XX 28-MAY-1999; 99US-0136782.

XX 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139839.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 26-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

12.5%; Score 69.5; DB 21; Length 334;

PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	16-AUG-1999;	99US-0148684.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160778.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Wed Jul 31 08:21:32 2002

PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137722.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 21-JUN-1999;	99US-0139763.
PR 22-JUN-1999;	99US-0139817.
PR 23-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140693.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 01-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142053.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143542.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 21-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.

PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 16-AUG-1999;	99US-0148684.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 22-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.
PR 23-OCT-1999;	99US-0161404.
PR 25-OCT-1999;	99US-0161405.
PR 25-OCT-1999;	99US-0161406.
PR 26-OCT-1999;	99US-0161359.
PR 26-OCT-1999;	99US-0161360.
PR 26-OCT-1999;	99US-0161361.
PR 28-OCT-1999;	99US-0161920.
PR 28-OCT-1999;	99US-0161992.
PR 28-OCT-1999;	99US-0161993.
PR 29-OCT-1999;	99US-0162142.

PR	27-MAY-1999;	9905-0136392.
PR	28-MAY-1999;	9905-0136782.
PR	30-MAY-1999;	9905-0137260.
PR	03-JUN-1999;	9905-0137528.
PR	04-JUN-1999;	9905-0137502.
PR	07-JUN-1999;	9905-0137724.
PR	08-JUN-1999;	9905-0138094.
PR	10-JUN-1999;	9905-0138540.
PR	10-JUN-1999;	9905-0138847.
PR	14-JUN-1999;	9905-0139119.
PR	16-JUN-1999;	9905-0139432.
PR	16-JUN-1999;	9905-0139453.
PR	17-JUN-1999;	9905-0139494.
PR	18-JUN-1999;	9905-0139455.
PR	18-JUN-1999;	9905-0139456.
PR	18-JUN-1999;	9905-0139457.
PR	18-JUN-1999;	9905-0139458.
PR	18-JUN-1999;	9905-0139459.
PR	18-JUN-1999;	9905-0139460.
PR	18-JUN-1999;	9905-0139461.
PR	18-JUN-1999;	9905-0139462.
PR	18-JUN-1999;	9905-0139463.
PR	18-JUN-1999;	9905-0139465.
PR	18-JUN-1999;	9905-0139750.
PR	18-JUN-1999;	9905-0139763.
PR	21-JUN-1999;	9905-0139817.
PR	22-JUN-1999;	9905-0139899.
PR	22-JUN-1999;	9905-0140353.
PR	23-JUN-1999;	9905-0140354.
PR	24-JUN-1999;	9905-0140695.
PR	24-JUN-1999;	9905-0140823.
PR	28-JUN-1999;	9905-0140991.
PR	29-JUN-1999;	9905-0141187.
PR	30-JUN-1999;	9905-0141842.
PR	01-JUL-1999;	9905-0142154.
PR	02-JUL-1999;	9905-0142055.
PR	06-JUL-1999;	9905-0142803.
PR	08-JUL-1999;	9905-0142820.
PR	09-JUL-1999;	9905-0142977.
PR	09-JUL-1999;	9905-0143342.
PR	12-JUL-1999;	9905-0143644.
PR	14-JUL-1999;	9905-0144005.
PR	15-JUL-1999;	9905-0144085.
PR	16-JUL-1999;	9905-0144086.
PR	19-JUL-1999;	9905-0144331.
PR	19-JUL-1999;	9905-0144332.
PR	19-JUL-1999;	9905-0144333.
PR	19-JUL-1999;	9905-0144334.
PR	19-JUL-1999;	9905-0144335.
PR	19-JUL-1999;	9905-0144332.
PR	20-JUL-1999;	9905-0144632.
PR	20-JUL-1999;	9905-0144884.
PR	21-JUL-1999;	9905-0144881.
PR	21-JUL-1999;	9905-0145066.
PR	21-JUL-1999;	9905-0145088.
PR	22-JUL-1999;	9905-0145085.
PR	22-JUL-1999;	9905-0145087.
PR	22-JUL-1999;	9905-0145089.
PR	22-JUL-1999;	9905-0145192.
PR	23-JUL-1999;	9905-0145145.
PR	23-JUL-1999;	9905-0145218.
PR	26-JUL-1999;	9905-0145276.
PR	27-JUL-1999;	9905-0145313.
PR	27-JUL-1999;	9905-0145918.
PR	28-JUL-1999;	9905-0145919.
PR	02-AUG-1999;	9905-0146581.
PR	02-AUG-1999;	9905-0146588.
PR	03-AUG-1999;	9905-0147038.

Query Match	12.2%	Score 68	DB 21	Length 322
Best Local Similarity	28.8%	Pred. No. 6.8	Mismatches 16	Indels 8
Matches 17	Conservative	18	Mismatches 16	Indels 8
3 NHHRLRTSALLT---	ALVYTGCVS--TGNYA---	MKEONOTIEORTIKRKTNRKOEIS	53	
: : :	: :	: : : :		
Db	58	nrvsvtsvaatteplvtlvtgagrgtqglvvykklketersqfvarglvrltkeesekin	116	
RESULT	9			
AAg1123				
ID	AAg1123	standard	Protein: 323 AA.	
XX	AAg1123			
AC	AAg1123			
XX	17-OCT-2000	(first entry)		
DT	17-OCT-2000			
XX	Arabidopsis thaliana	protein fragment	SEQ ID NO: 9722.	
DE	Arabidopsis thaliana	protein fragment	SEQ ID NO: 9722.	
XX	Protein identification: signal transduction pathway; metabolic pathway; hydri	disatation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.			
KW	Arabidopsis thaliana.			
XX	EP1033405-A2.			
PN	EP1033405-A2.			
XX	06-SEP-2000.			
PD	06-SEP-2000.			
XX	25-FEB-2000; 2000EP-0301439.			
PF	25-FEB-2000; 2000EP-0301439.			
XX	25-FEB-1999; 99US-0121825.			
PR	05-MAR-1999; 99US-0123180.			
PR	09-MAR-1999; 99US-0123548.			
PR	23-MAR-1999; 99US-0125788.			
PR	25-MAR-1999; 99US-0126264.			
PR	29-MAR-1999; 99US-0126785.			
PR	01-APR-1999; 99US-0127462.			
PR	06-APR-1999; 99US-0128234.			
PR	08-APR-1999; 99US-0128714.			
PR	16-APR-1999; 99US-0129845.			
PR	19-APR-1999; 99US-0130077.			
PR	23-APR-1999; 99US-0130449.			
PR	23-APR-1999; 99US-0130510.			
PR	23-APR-1999; 99US-0130891.			
PR	28-APR-1999; 99US-0131449.			
PR	30-APR-1999; 99US-0132048.			
PR	30-APR-1999; 99US-0132407.			
PR	04-MAY-1999; 99US-0132484.			
PR	05-MAY-1999; 99US-0132485.			
PR	06-MAY-1999; 99US-0132486.			
PR	06-MAY-1999; 99US-0132487.			
PR	07-MAY-1999; 99US-0132863.			
PR	11-MAY-1999; 99US-0134256.			
PR	14-MAY-1999; 99US-0134218.			
PR	14-MAY-1999; 99US-0134219.			
PR	14-MAY-1999; 99US-0134221.			
PR	14-MAY-1999; 99US-0134370.			
PR	18-MAY-1999; 99US-0134768.			
PR	19-MAY-1999; 99US-0134941.			
PR	20-MAY-1999; 99US-0135124.			
PR	21-MAY-1999; 99US-0135353.			
PR	24-MAY-1999; 99US-0135629.			
PR	25-MAY-1999; 99US-0136021.			
PR	27-MAY-1999; 99US-0136392.			
PR	28-MAY-1999; 99US-0136782.			
PR	01-JUN-1999; 99US-0137222.			
PR	03-JUN-1999; 99US-0137528.			
PR	04-JUN-1999; 99US-0137502.			
PR	07-JUN-1999; 99US-0137724.			
PR	08-JUN-1999; 99US-0138094.			

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139765.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 29-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142055.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144336.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148564.
PR 13-AUG-1999; 99US-0149368.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157865.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0160741.
PR 18-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 12.2%; Score 68; DB 21; Length 323;
Best Local Similarity 28.8%; Pred. No. 6.9';
Matches 17; Conservative 18; Mismatches 16; Indels 8; Gaps 3;
OY 3 NHHIRITISALT---ALLVTGCVS-TGNVA--*KEDNOQRIEFTIIKGTNKNQELS 53

Db 59 nrtsvtsaaatleplvltvgaagrtgqlyvkkikerseqfvarglvrtkeskekin 117
RESULT 10
AAG1122
ID AAG1122 standard; Protein; 325 AA.
AC AAG1122;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9721.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydrolisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123348.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

```

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158393.
PR 13-OCT-1999; 99US-0158993.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160776.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 12.2%; Score 68; DB 21; Length 325;

Best Local Similarity 28.8%; Pred. No. 6.9;

Matches 17; Conservative 18; Mismatches 16; Indels 8; Gaps 3;

```

OY 3 NHHIRLTIALLT-----ALIVTGVSTGNVNA---MKRONOQITOTIKTKTKRKQETS 53
    | :|:| | | :|:| | | | :|:|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 61 nrvsvtvsaaatpvtvltgaggrtgivvykkikrsefvarglvrtkkeskekin 119

```

RESULT 11
AAR95910.
ID AAR95910 standard; Protein; 151 AA.

```

XX AAR95910;
AC 17-AUG-1996 (first entry)
XX Opacity associated protein Oapb.
DE Opacity associated protein; OapA; OapB; vector; vaccine; antibody;
XX diagnosis; immunogen.
XX Haemophilus influenzae strain Rd variant H175.
OS W09613515-A1.
XX 09-MAY-1996.
PD 24-OCT-1995; 95WO-US13672.
XX 31-OCT-1994; 94US-0332576.
XX (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
PA Weiser JN;
XX WPI; 1996-239445/24.
XX N-PSDB; AAT27729.
DR New opacity associated protein of Haemophilus influenzae - and
PT related DNA, vectors, transformed hosts and antibodies, useful in
PT vaccines and for diagnosis.
XX Claim 3; Fig 1a-c; 48pp; English.
XX Haemophilus influenzae opacity associated proteins OapA (AAR95909)
CC and OapB (AAR95910) are membrane proteins useful as immunogens in
CC and OapB (AAR95910) are membrane proteins useful as immunogens in
CC vaccines against typhoid and non-typhoid H. influenzae infection.
CC They were identified as the products of 2 open reading frames of
CC an oap gene (AAT27729) isolated from the spontaneous opacity variant
CC H175. Recombinant Oap proteins can be expressed in host cells
CC carrying a vector including an oap gene sequence. Antibodies
CC raised against the proteins are useful for detecting H. influenzae
CC or for passive immunisation.
XX Sequence 151 AA;
SQ

```

Query Match 12.1%; Score 67.5; DB 17; Length 151;
Best Local Similarity 33.3%; Pred. No. 2.9; Mismatches 15; Indels 3; Gaps 1;
Matches 14; Conservative 10;

```

OY 11 SALLTALLVTGCVSTGNVAMKRONOQITIE--QTIIKGTNK 49
    | :|:| | | | | | :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 6 sltfallmtgcvgvnanvtpgagkmgvkvdkalgysedr 47

```

RESULT 12

ABG02873 standard; Protein; 222 AA.

ABG02873;

13-FEB-2002 (first entry)

Novel human diagnostic protein #2864.

Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

WO200175067-A2.

11-OCT-2001.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/895,300A
 FILING DATE: 19920608
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Hensley, Max D.
 REGISTRATION NUMBER: 27,043
 REFERENCE/DOCKET NUMBER: 747
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-1994
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 346 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear

[illegible]

APPLICATION NUMBER: 08/116186
FILING DATE: 02-Sep-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/895300
FILING DATE: 08-Jun-1992
ATTORNEY/AGENT INFORMATION:
NAME: Johnston, Sean A.
REGISTRATION NUMBER: 35,910
REFERENCE/DOCKET NUMBER: P0747C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3562
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-458-367-1

Query Match 11.4%; Score 63.5; DB 1; Length 346;
Best Local Similarity 24.2%; Pred. No. 3.4;
Matches 30; Conservative 23; Mismatches 42; Indels 29; Gaps 6;

OY 3 NHHILRTISALTALVYGVSTGVNANK--EONOQTEOTIKGTNKQ---EISSRFG 57
:|:|:| | | | | :|:|:| | | :|:|:| | | :|:|:| | | :|:|:| | | :|:|:| | |
Db 49 HHHLRMGMKLLGALLAALQAVSLKIAFNQTGETMSNATLVSYIVQLSRD 108
OY 58 SADSIEMIVIKFGHTAIL-----APNRMOEILSLIIS-----FLWVKPYRP 100
:|:|:| | | | | :|:|:| | | :|:|:| | | :|:|:| | | :|:|:| | | :|:|:| | |
Db 109 IA-----LVQEVDSHLTAVGKLDLNLQDAPDQTYHYVVSPLGHSYKERYLFV--YRP 161
OY 101 KTLS 104
:|
Db 162 DOVS 165

RESULT 5
PCT-US93-05136-1
Sequence 1, Application PC/TUS9305136
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: PURIFIED FORMS OF DNase
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: palin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/05136
FILING DATE: 19930528
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Johnston, Sean A.
REGISTRATION NUMBER: P35,910
REFERENCE/DOCKET NUMBER: 747PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3562
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US93-05136-1

Query Match 11.4%; Score 63.5; DB 5; Length 346;
Best Local Similarity 24.2%; Pred. No. 3.4;
Matches 30; Conservative 23; Mismatches 42; Indels 29; Gaps 6;

OY 3 NHHILRTISALTALVYGVSTGVNANK--EONOQTEOTIKGTNKQ---EISSRFG 57
:|:|:| | | | | :|:|:| | | :|:|:| | | :|:|:| | | :|:|:| | | :|:|:| | |
Db 49 HHHLRMGMKLLGALLAALQAVSLKIAFNQTGETMSNATLVSYIVQLSRD 108
OY 58 SADSIEMIVIKFGHTAIL-----APNRMOEILSLIIS-----FLWVKPYRP 100
:|:|:| | | | | :|:|:| | | :|:|:| | | :|:|:| | | :|:|:| | | :|:|:| | |
Db 109 IA-----LVQEVDSHLTAVGKLDLNLQDAPDQTYHYVVSPLGHSYKERYLFV--YRP 161
OY 101 KTLS 104
:|
Db 162 DOVS 165

RESULT 6
US-08-651-999A-3
Sequence 3, Application US/08651999A
Patent No. 6031088

GENERAL INFORMATION:
APPLICANT: Stefan Somlo and Toshio Mochizuki
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE PKD2 GENE
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
STREET: 90 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE
MEDIUM TYPE: DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,999A
FILING DATE: MAY 23, 1996
ATTORNEY/AGENT INFORMATION:
NAME: ELIZABETH A. BOGOSTIAN
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 96700/395
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
DESCRIPTION: NO
HYPOTHETICAL: NO
FRAGMENT TYPE: internal fragment
US-08-651-999A-3

Query Match 10.7%; Score 59.5; DB 3; Length 399;
Best Local Similarity 23.9%; Pred. No. 14;
Matches 26; Conservative 14; Mismatches 40; Indels 29; Gaps 3;

APPLICANT: Deretic, Vojo
APPLICANT: Martin, Daniel W.
TITLE OF INVENTION: DETECTION OF CONVERSION TO MUCOID IN
TITLE OF INVENTION: PSEUDOMONAS AERUGINOSA INFECTING CYSTIC FIBROSIS PATIENTS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/505,307
FILING DATE: 24-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/017,114
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hodgins, Daniel S.
REGISTRATION NUMBER: 31,026
REFERENCE/DOCKET NUMBER: UTSK:231
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-505-307-19

Query Match 10.28; Score 57; DB 3; Length 193;
Best Local Similarly 23.28; Pred. No. 11;
Matches 19; Conservative 13; Mismatches 12; Indels 38; Gaps 3;
QY 31 KEONQITIEGTTIKGKTKNKQISRFSGSADISIFMIVYIKFGHTAILAPNRMQETLSLI 90
DB 4 OEODQOLVE-----RVQRGDKRAFDLVLTQTH-----KILGLIV 38
QY 91 SFL-----WVKPYR 99
DB 39 RFVHDQAEADVDVAQAEAFIKAYR 60

Search completed: July 30, 2002, 15:38:55
Job time: 184 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:37:11 ; Search time 16.29 seconds
(without alignments)
654.752 Million cell updates/sec

Title: US-10-048-197-2

Perfect score: 558
Sequence: 1 MLNHHIRLTISALLTALLVT.....FLWVKPRPKNLSPYLTKAKA 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117.5	21.1	131	2 B91067	hypothetical lipop
2	78	14.0	289	2 H95216	cation efflux syst
3	78	14.0	299	2 F98080	hypothetical prote
4	74	13.3	328	2 AH2128	iron(III) dicitrat
5	72.5	13.0	112	2 B81969	hypothetical prote
6	72.5	13.0	138	2 S74868	hypothetical prote
7	71.5	12.8	226	2 C70078	hypothetical prote
8	71	12.7	568	2 T49962	hypothetical prote
9	70.5	12.6	216	2 T40178	60s ribosomal prot
10	70	12.5	518	2 E71653	protein-export mem
11	69.5	12.5	314	1 G64145	hypothetical prote
12	69.5	12.5	337	2 T02532	hypothetical prote
13	69	12.4	530	2 T29418	hypothetical prote
14	68.5	12.3	391	2 C71499	probable gen. secr
15	68.5	12.3	405	2 E71179	hypothetical prote
16	68	12.2	430	2 E86873	proteinase (import
17	67.5	12.1	134	2 B64062	opacity-associated
18	67.5	12.1	210	2 E83816	late competence op
19	67.5	12.1	216	2 T40848	60s ribosomal prot
20	67	12.0	455	1 T21089	acid phosphatase (
21	67	12.0	488	1 H64333	corrinoid/iron-sul
22	67	12.0	799	2 B86735	hypothetical prote
23	67	12.0	819	2 G69801	hypothetical prote
24	66.5	11.9	92	2 B85911	hypothetical prote
25	66.5	11.9	310	2 G97345	oligopeptide ABC t
26	66.5	11.9	364	2 S48121	transcription fact
27	66.5	11.9	364	2 T50007	transcription fact
28	66.5	11.9	429	2 G71937	hypothetical prote
29	66	11.8	396	2 T01049	hypothetical prote

30	66	11.8	458	2 F71698	hypothetical prote
31	66	11.8	599	2 E36792	hypothetical prote
32	66	11.8	1828	2 B59254	myosin heavy chain
33	66	11.8	1855	2 A59254	myosin heavy chain
34	65.5	11.7	103	2 T58936	hypothetical prote
35	65.5	11.7	161	2 T08739	hypothetical prote
36	65.5	11.7	182	2 G69305	hypothetical prote
37	65.5	11.7	319	2 S62196	hypothetical prote
38	65.5	11.7	345	2 F90239	anthranilate phosp
39	65.5	11.7	419	2 B95106	conserved hypotet
40	65.5	11.7	425	2 H82183	conserved hypotet
41	65.5	11.7	629	2 T18227	hypothetical prote
42	65.5	11.7	690	2 H71237	hypothetical prote
43	64.5	11.6	196	2 G72423	hypothetical prote
44	64.5	11.6	419	2 D97974	conserved hypotet
45	64.5	11.6	492	2 T06875	preprotein translo

ALIGNMENTS

```

RESULT 1
B91067
hypothetical lipoprotein [imported] - Escherichia coli (strain O157:H7, substrain R1M
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: B91067
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A96629; MUID:21156231; PMID:11258796
A:Accession: B91067
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-131 <HAV>
A:Cross-references: GB:BA000007; PIDN:BAR36929.1; PID:G13362977; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1M 0509952
C:Genetics:
A:Gene: ECs3506

Query Match 21.1%; Score 117.5; DB 2; Length 131;
Best Local Similarity 29.7%; Pred. No. 4.3e-05;
Matches 27; Conservative 19; Mismatches 30; Indels 15; Gaps 1;

Oy 17 LVTGCVSTGVAMKEONQOTIEOTIIKGTNKQKDISRFGSADISFMIVIRFGHT-- 74
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 19 LTLGCSVGNQSLKNETQESVTKRIYKGTTKODVLASFGEPSRSRLIDEEQMSYTMV 78
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 75 -----AIIAPNRQGLSLISF 92
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 79 NSQSKATSFIPVGLAGGADSDTKSLTVSF 109

RESULT 2
H95216
cation efflux system protein [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: H95216
R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappel
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: H95216
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK75929.1; PID:G14973359; GSPDB:GN00164; TIGR:

```


A: Experimental source: strain Madrid E
 C: Genes: RP586
 A: Gene: secD; protein export membrane protein secD
 C: Superfamily: protein export membrane protein secD

Query Match 12.5%; Score 70; DB 2; Length 518;
 Best Local Similarity 26.5%; Pred. No. 19;
 Matches 31; Conservative 21; Mismatches 39; Indels 26; Gaps 6;
 Oy 8 LTTSLATLALVTCVSTGVAMKEONQOTIEOTIKGKTNRKOEISSRFGSA-----DS 61
 Db 404 LTPGILGILITIGMAVDANVLYER-----IKRELKGVSNLYAIKTFESAFATIDSN 459
 Oy 62 ISPMIVIVIK---FGHTAI-----IAPNRWQEL--SLTISLWKPYPKRL 103
 Db 460 ITTLIVAFALYIFGVGAIKGFVALFTIGLISMSFSAIITRKLLID-VWVKYFAPKRL 515

RESULT 11

G64145
 hypothetical protein H10230 - Haemophilus influenzae (strain Rd KM20)
 C: Species: Haemophilus influenzae
 C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jul-2000
 C: Accession: G64145
 R: Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kervatage, J.; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M. Science 269, 496-512, 1995
 A: Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A. Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A: Reference number: A64000; MUID: 95350630
 A: Accession: G64145
 A: Status: nucleic acid sequence not shown; translation not shown
 A: Molecule type: DNA
 A: Residues: 1-314 <TIGR>
 A: References: GB:U02709; GB:I42023; NID:g1573190; PIDN:AC021899.1; PID:g1573194; T
 A: Note: best homolog was a hypothetical protein from Escherichia coli
 C: Superfamily: hypothetical protein H10230; tetratricopeptide repeat homology <TT1>
 F: 74-107/Domain: tetratricopeptide repeat homology <TT2>
 F: 108-141/Domain: tetratricopeptide repeat homology <TT3>
 F: 142-175/Domain: tetratricopeptide repeat homology <TT3>

Query Match 12.5%; Score 69.5; DB 1; Length 314;
 Best Local Similarity 27.1%; Pred. No. 12;
 Matches 19; Conservative 13; Mismatches 21; Indels 17; Gaps 2;
 Oy 2 LNHRIITLITLALVTCG-----VSTGVAMKEONQOT-----IEGTIK 44
 Db 6 LSRHFIYLFSLCALILLACVSRGFGVSKNHNVLAEQNPNTHFQEVMIYRLSQLLV 65
 Oy 45 GKTNRKOEISS 54
 Db 66 GKNSNERAS 75

RESULT 12

T02532
 hypothetical protein At2g37660 [imported] - Arabidopsis thaliana
 N: Alternate names: hypothetical protein F13M22.16
 C: Species: Arabidopsis thaliana (mouse-ear cross)
 C: Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
 C: Accession: T02532; E84795
 R: Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, June 1998
 A: Description: Arabidopsis thaliana chromosome II BAC F13M22 genomic sequence.
 A: Reference number: Z14677
 A: Accession: T02532
 A: Status: translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 1-337 <ROU>
 A: Cross-references: EMBL:AC004684; NID:g3236234; PID:g3236248

A: Experimental source: cultivar Columbia
 R: Lin, X.; Kaul, S.; Rounsley, S.D.; Shue, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, M.; Koo, H.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999
 A: Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A: Reference number: A84420; MUID: 20083487
 A: Accession: E84795
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-337 <STO>
 A: Cross-references: GB:AE002093; NID:g3236248; PIDN:AC023636.1; GSPDB:GN00139

Query Match 12.5%; Score 69.5; DB 2; Length 337;
 Best Local Similarity 23.2%; Pred. No. 13;
 Matches 36; Conservative 26; Mismatches 38; Indels 55; Gaps 8;
 Oy 3 NHHIRLITLALVTCGVS--TGNYA---MKRONQOTIEOTIKGKTNRKOEISS 54
 Db 61 NRRVSVTVSAATTEPLVLTGAGRGITGVYKKLEBSEQFVABGLVTKESKEKING 120
 Oy 55 RGSADISPMIVIVIKFGHTAILAPNRWQELSLTISLWKPYPKRL----- 100
 Db 121 E-----VFQIDIDFTASIAV-AVEGIDALVILTSAPVPMKPGFDPKSGRPEFF 171
 Oy 101 -----KN-----ISFYLTANA 111
 Db 172 DGAAPVEQWIGOKNOIDAGVDDGLSFVYVAKA 206

RESULT 13

T29418
 hypothetical protein F21F8.11 - Caenorhabditis elegans
 C: Species: Caenorhabditis elegans
 C: Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C: Accession: T29418
 R: Wilson, R.; Favello, A.; Le, T.T.
 A: Submitted to the EMBL Data Library, April 1997
 A: Description: The sequence of C. elegans cosmid F21F8.
 A: Reference number: Z20618
 A: Accession: T29418
 A: Status: preliminary; translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 1-530 <MIL>
 A: Cross-references: EMBL:U97000; PIDN:AC047997.1; GSPDB:GN00023; CESP:F21F8.11
 A: Experimental source: strain Bristol N2; clone F21F8
 C: Genes: CESP:F21F8.11
 A: Map position: 5
 A: Introns: 16/1; 64/2; 112/3; 344/3; 417/1; 452/1

Query Match 12.4%; Score 69; DB 2; Length 530;
 Best Local Similarity 21.7%; Pred. No. 24;
 Matches 30; Conservative 25; Mismatches 47; Indels 36; Gaps 5;

Oy 5 HIRLITLALVTCVSTGVAMKEONQOTIEOTIKGKTNRKOEISSRFGSAD---- 60
 Db 44 HIGLSLSCMCNSTAV-ALMNTNNTLVEGSTILMSLENKTSSE---SQELGGDDGES 99
 Oy 61 -----SISPMIVIVIK-----FGHTAILAPNRW---QELISLISF 92
 Db 100 CTKLESKYIKYDGGFINSVSMOGIYSAFLGFIFFYPAGVLDVRSARHILSVAILM 159
 Oy 93 LWKPYRPNKLSFYLTAK 110
 Db 160 LTIASLMPVLISYIGEK 177

RESULT 14

C71499

probable gen. secretion protein F - Chlamydia trachomatis (serotype D, strain UW3/Cx)

C:Species: Chlamydia trachomatis

C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 17-Mar-2000

C:Accession: C71499

R:Stephens, R.S.; Kaiman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis

A:Reference number: A71570; MUID:9900809

A:Accession: C71499

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-391 <ARN>

A:Cross-references: GB:AE00327; GB:AE001273; NID:93328999; PIDN:AC68172.1; PID:9332901

A:Experimental source: serotype D, strain UW-3/Cx

C:Genetics:

A:Gene: g9pF

C:Superfamily: secretion protein xcps

Query Match

12.3%; Score 68.5; DB 2; Length 391;

Best Local Similarity 21.7%; Pred. No. 20;

Matches 20; Conservative 22; Mismatches 39; Indels 11; Gaps 3;

OY 14 LTALLVGCSTGNVAMKEONQOTIEQTIKGTNKOESRFGSADSISFMI----- 66

DB 275 LIEALTYLGCFAVSODFLREELQEVY-QAVVKGSLSRSLSHRTWTPKLVIGVALGESG 333

OY 67 -VVIKFGHTAILAPNRMOEILSLISFLWKP 97

DB 334 DLAVFAHVAQIYVEDIORVLTWTA--WCQP 363

RESULT 15

E71179

hypothetical protein PH1714 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000

C:Accession: E71179

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekin, M.; Ohtuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon: Pyrococcus horikoshii

A:Reference number: A71000; MUID:98344137

A:Accession: E71179

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-405 <RAW>

A:Cross-references: GB:AP000007; NID:93236134; PIDN:BA30828.1; PID:93258145

A:Experimental source: strain OT3

C:Genetics:

A:Gene: PH1714

Query Match

12.3%; Score 68.5; DB 2; Length 405;

Best Local Similarity 22.5%; Pred. No. 20;

Matches 27; Conservative 17; Mismatches 33; Indels 43; Gaps 4;

OY 6 IRLTISALTLVLTGCVSTGNVAMKEONQOTIEQTIKGTNKOESRFGSADSISF- 64

DB 5 LALSTIGLVLSVAGCTGGG-----TOTOTOGKSIVAILFDVGGRGDLSFN 54

OY 65 -----MIVVIFGHTAILAPNRMOEILSL-----IISFLWKP 97

DB 55 DMAVYGAERAKKELGVKIEY-----MTPKSKEDVPLLEQLAKSKEYDLVLVGFLLWTP 109

Search completed: July 30, 2002, 15:39:18
Job time: 127 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:38:56 ; Search time 11.9 Seconds
(without alignments)
361.166 Million cell updates/sec

Title: US-10-048-197-2
Perfect score: 558
Sequence: 1 MLNHHIRLTALTLALT.....FLWVKPYRPNLSFYLTAKA 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	71.5	12.8	226 1	YXIP_BACSU
2	70.5	12.6	216 1	R10A_SCHPO
3	70	12.5	518 1	SECD_RICPR
4	69.5	12.5	314 1	NLPI_HAEIN
5	68.5	12.5	337 1	Y230_ARATH
6	68.5	12.3	405 1	YH14_PYRHO
7	67.5	12.1	134 1	OAPB_HAEIN
8	67.5	12.1	216 1	R10B_SCHPO
9	66	11.8	370 1	RF1_MYCCA
10	66	11.8	599 1	VG58_HSV11
11	66	11.8	1855 1	MY5A_HUMAN
12	65.5	11.7	345 1	TRPD_SUISO
13	64.5	11.6	492 1	SECT_CYAPA
14	64	11.5	157 1	Y406_MYCPN
15	64	11.5	880 1	RPAL_SUISO
16	64	11.5	1828 1	MY5A_RAT
17	64	11.5	1829 1	MY5A_CHICK
18	64	11.5	1853 1	MY5A_MOUSE
19	63.5	11.4	986 1	DROL_NPYBM
20	63	11.3	254 1	T2H1_HAEPA
21	63	11.3	262 1	YF10_HAEIN
22	62.5	11.2	984 1	DROL_NPYAC
23	62.5	11.2	1145 1	PR22_YEAST
24	62.5	11.2	3206 1	POLG_PSBMV
25	62	11.1	433 1	ORC4_MOUSE
26	62	11.1	442 1	YMD8_YEAST
27	62	11.1	1062 1	YAF1_YEAST
28	61.5	11.0	467 1	LOLS_DROME
29	61.5	11.0	894 1	LOLL_DROME
30	61	10.9	265 1	SYNP_RAT
31	61	10.9	288 1	UBIA_PROST
32	61	10.9	1940 1	MYH3_HUMAN
33	61	10.9	1940 1	MYH3_RAT

34	60.5	10.8	201 1	OSMY_ECOLI
35	60	10.8	335 1	FLIG_THEMA
36	60	10.8	359 1	MCAL_CRIGR
37	60	10.8	367 1	YC36_HAEIN
38	60	10.8	475 1	NCAP_STNV
39	60	10.8	571 1	TACY_STRPY
40	60	10.8	603 1	NU5M_RABIT
41	60	10.8	654 1	RPC3_YEAST
42	60	10.8	1091 1	SYL_TREPA
43	59.5	10.7	330 1	EYL2_HUMAN
44	59.5	10.7	414 1	YBBC_BACSU
45	59.5	10.7	568 1	MET3_ASPTPE

ALIGNMENTS

RESULT 1	YXIP_BACSU	STANDARD:	PRT:	226 AA.
AC	P42307:			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Hypothetical protein yxip precursor.			
GN	YXIP OR S3B.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Bacillus.			
OX	NCBI_TaxID=1423;			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=168 / BGSC1A1;			
RC	MEDLINE=97124196; PubMed=8969509;			
RX	Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,			
RA	Miwa Y., Fujita Y.;			
RT	"Sequencing of a 65 kb region of the Bacillus subtilis genome			
RT	containing the ltc and cel loci, and creation of a 177 kb contig			
RT	covering the gnt-sacXY region."			
RL	Microbiology 142:3113-3123(1996).			
RN	[2]			
RP	SEQUENCE OF 176-226 FROM N.A.			
RC	STRAIN=BR151;			
RA	Stuelse J., Schnetz K., Krieg M., Krueger S., Hecker M., Rak B.;			
RL	Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; D83026; BAAL1695.1; -			
DR	EMBL; 228340; -; NOT_ANNOTATED_CDS.			
DR	EMBL; 299124; CAB15945.1; -			
DR	Subtilist; BC11145; yxip.			
DR	PROSITE; PS00013; PROKAR_LIPOPROTEIN; UNKNOWN_1.			
RW	Hypothetical protein; Signal; Complete proteome.			
FT	SIGNAL 1 18			POTENTIAL.
FT	CHAIN 19 226			HYPOTHETICAL PROTEIN YXIP.
SQ	SEQUENCE 226 AA; 25697 MW; 49D7A86DF4400D08 CRC64;			

Query Match 12.8%; Score 71.5; DB 1; Length 226;
Best Local Similarity 36.2%; Pred. No. 1.9;
Matches 17; Conservative 13; Mismatches 16; Indels 1; Gaps 1;

OY 6 IRLTISALTALVTGCVSTGVNVA-KRONOOTIOTIITKGTNKE 51
DB 4 IGLTISLTVTIVMSACESEGEAOMFADCDKOTVAKPSSKK 50

```

RESULT 2
ID R10A_SCHPO STANDARD: PRT: 216 AA.
AC 014363;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S ribosomal protein L1-A (L10A).
GN RPLA OR SPBC30D10.18C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_Taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Rajandream M.A., Barrell B.G., Lauber J., Hilbert H.,
RA Duesterhoeft A.; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L1 IN S. POMBE.
CC -1- SIMILARITY: BELONGS TO THE L1P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 297992; CAB10813.1;
CC InterPro: IPR002143; Ribosomal_L1.
DR Pfam: PF00687; Ribosomal_L1.
DR ProDom: PD001314; Ribosomal_L1.
DR PROSITE: PS01199; RIBOSOMAL_L1.
KW Ribosomal protein; Multigene family.
SQ SEQUENCE 216 AA; 23884 MW; 676EA3710A497165 CRC64;

Query Match 12.6%; Score 70.5; DB 1; Length 216;
Best Local Similarity 26.6%; Pred. No. 2.4;
Matches 25; Conservative 16; Mismatches 24; Indels 29; Gaps 4;

OY 6 IRLTISALTLALVTGCVSTGVNAAKEQNOQTETIIRKTKNKOISRFSGADSIFM 65
DB 150 VKSTIKPOLKVLCLG-VAVGHVDMAB-----EQLANLSLA--INFL 189
OY 66 IVIKFEGHTAIAPNRMOELSLIISFLWKKPYR 99
DB 190 VSLRLKG-----WONIGSLVIKSTMGKPRR 214

RESULT 3
ID R10A_SCHPO STANDARD: PRT: 518 AA.
AC 092CW8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein-export membrane protein secD.
GN SEC2 OR RP586.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsia.
OX NCBI_Taxid=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RX Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RX Sichenitz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,

```

```

RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."
RL Nature 396:133-140(1998).
CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT.
CC -1- SUBUNIT: PART OF THE PROKARYOTIC PROTEIN TRANSLLOCATION APPARATUS
CC WHICH COMPRISE SECA, SECB, SECD, SECE, SECF, SECG AND SECH
CC (BY SIMILARITY)
CC -1- SIMILARITY: BELONGS TO THE SEC2/SECF FAMILY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ235272; CA15031.1;
CC InterPro: IPR003335; SecD_SecF.
DR Pfam: PF02355; SecD_SecF.
DR Protein transport; Translocation; Transmembrane; Membrane;
KW Complete proteome.
FT TRANSMEM 9 POTENTIAL.
FT TRANSMEM 356 376 POTENTIAL.
FT TRANSMEM 377 397 POTENTIAL.
FT TRANSMEM 406 426 POTENTIAL.
FT TRANSMEM 464 484 POTENTIAL.
FT TRANSMEM 486 506 POTENTIAL.
SQ SEQUENCE 518 AA; 56712 MW; 86D6FC2A5B5DE2AA CRC64;

Query Match 12.5%; Score 70; DB 1; Length 518;
Best Local Similarity 26.5%; Pred. No. 6.9;
Matches 31; Conservative 21; Mismatches 39; Indels 26; Gaps 6;

OY 8 LRTISALTLALVTGCVSTGVNAAKEQNOQTETIIRKTKNKOISRFSGA-----DS 61
DB 404 LTLPGIAGILITIGMAVDANVLIER-----IKEELHKVSNLSAIRTFESAFATIDSN 459
OY 62 ISEFIVIRK---FGHTAI-----LAPNRMOEL--SLIISFLWKKPYRPNL 103
DB 460 ITTLIVAFALYIGVAGIKGFVAVALTIGIISMFSAITTKLLID-VWYKPKPKKL 515

RESULT 4
ID NLP1_HAEIN STANDARD: PRT: 314 AA.
AC P44585;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lipoprotein nlp1 homolog precursor.
GN NLP1 OR HI0230.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_Taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RX Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RX McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RX Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RX Weidman J.F., Phillips R., Liu L.-I., Spriggs T., Hedblom E., Cotton M.D.,
RX Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RX Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RX Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus

```



```

RT Influenzae Rd.*
RL Science 269:496-512(1995).
CC -1- FUNCTION: MAY BE INVOLVED IN CELL DIVISION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (By similarity).
CC -1- SIMILARITY: CONTAINS 3 TPR REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U33709; AAC21899.1; -.
DR TIGR: H10230; -.
DR InterPro: IPR001440; TPR.
DR Pfam: PF00515; TPR; 3.
KW Cell division; Membrane; Lipoprotein; Repeat; TPR repeat; Signal;
KM Complete proteome.
FT SIGNAL 1 25 BY SIMILARITY.
FT CHAIN 26 314 LIPOPROTEIN NLPT HOMOLOG.
FT LIPID 26 26 N-ACYL DIGLYCERIDE.
FT REPEAT 74 107 TPR 1.
FT REPEAT 108 141 TPR 2.
FT REPEAT 142 175 TPR 3.
SQ SEQUENCE 314 AA; 3643 MW; 0B87CAE286CC5722 CRC64;

```

```

Query Match 12.5%; Score 69.5; DB 1; Length 314;
Best Local Similarity 27.1%; Pred. No. 4.5;
Matches 19; Conservative 13; Mismatches 21; Indels 17; Gaps 2;

```

```

QY 2 LNHRIRLTISALTLALVTGCG-----VSTGNVAMKEQNOOT-----IEQRTIK 44
   1 : : : : : 11 : : : : : 1 : : : : :
DB 6 LSRHFIVYLFSCAILLILGCVQSRGFGVSKNHVYLAENQNTHEQDEMTIVRLSQVLLV 65
   1 : : : : : 1 : : : : : 1 : : : : :
QY 45 GKTNKOEISS 54
   1 : : : : : 1 : : : : : 1 : : : : :
DB 66 GKMSNEERAS 75

```

```

RESULT 5
Y230_ARATH STANDARD: PRT; 337 AA.
ID Y230_ARATH
AC 080934;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein At2g37520, chloroplast precursor.
GN AT2G37520 OR F13M22.16.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Motilal K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Mearns W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RL Nature 402:761-768(1999).
CC -1- SUBCELLULAR LOCATION: Chloroplast (Potential).

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AC004684; AAC23636.1; -.
DR SWISS-2DPAGE: 080934; ARATH.
KW Chloroplast; Transil peptide.
FT TRANSIT 1 337 CHLOROPLAST (POTENTIAL).
FT CHAIN ? 337 PROTEIN AT2G37520.
SQ SEQUENCE 337 AA; 36085 MW; 748838ECB4BA07CD CRC64;

```

```

Query Match 12.5%; Score 69.5; DB 1; Length 337;
Best Local Similarity 23.2%; Pred. No. 4.9;
Matches 36; Conservative 26; Mismatches 38; Indels 55; Gaps 8;

```

```

QY 3 NHHRIRLTISALT-----ALVTGCVS--TGNVA---MKEQNOOTIEQTIKTKTKOEISS 54
   1 : : : : : 1 : : : : : 1 : : : : : 1 : : : : :
DB 61 NRRSVTVSAATTEPLVTLVTGAGGRTGQIVYKRLKRSQFVARKGLVTRKSEKING 120
   1 : : : : : 1 : : : : : 1 : : : : : 1 : : : : :
QY 55 RFGSADSIEMIVIKFGHTAILAPNRMOELISLISFLWKPYRP----- 100
   1 : : : : : 1 : : : : : 1 : : : : : 1 : : : : :
DB 121 E-----DE-----VFIGDIRDTASINP-AVEGIDALVILTSANVPQMKFGDPKSGARPEFF 171
   1 : : : : : 1 : : : : : 1 : : : : : 1 : : : : :
QY 101 -----KN-----LSFYLAKA 111
   1 : : : : : 1 : : : : : 1 : : : : : 1 : : : : :
DB 172 DDGAVPEQVDWIGQKNQIDAGVDDGSLFVYTAKA 206

```

```

RESULT 6
YH14_PYRHO STANDARD: PRT; 405 AA.
ID YH14_PYRHO
AC 059403;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical Lipoprotein PH1714 precursor.
GN PH1714 OR PHAM037.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohnuki Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Maechli Y., Shizuwa H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000007; BAA30828.1; -.
DR InterPro: IPR003760; Bmp.

```


RP SEQUENCE OF 1358-1460 FROM N.A. (ISOFORM 3).
 RA Edgar A.J., Bennett J.P.;
 RT "Inhibition of dendrite formation in melanocytes transiently
 RT transacted with antisense DNA to myosin V.";
 RT Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
 RL [7]
 RN FUNCTION.
 RP MEDLINE=99376094; PubMed=10448864;
 RX Mehta A.D., Rock R.S., Rief M., Spudich J.A., Mooseker M.S.,
 RA Cheney R.E.;
 RT "Myosin-V is a processive actin-based motor."
 RT Nature 400:590-593(1999).
 CC -1- FUNCTION: PROCESSIVE ACTIN-BASED MOTOR THAT CAN MOVE IN LARGE
 CC STEPS APPROXIMATING THE 36-NM PSEUDO-REPEAT OF THE ACTIN FILAMENT.
 CC MAY BE INVOLVED IN MELANOSOME TRANSPORT, OR ALTERNATIVELY, IT MAY
 CC BE REQUIRED FOR SOME POLARIZATION PROCESS INVOLVED IN DENDRITE
 CC FORMATION.
 CC -1- SUBUNIT: MAY BE A HOMODIMER, WHICH ASSOCIATES WITH MULTIPLE
 CC CALMODULIN OR MYOSIN LIGHT CHAINS.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- DISEASE: DEFECTS IN MYOSA ARE A CAUSE OF GRISCELLI SYNDROME (GS).
 CC GS IS A RARE AUTOSOMAL RECESSIVE DISORDER THAT RESULTS IN
 CC PIGMENTARY DILUTION OF THE SKIN AND HAIR, THE PRESENCE OF LARGE
 CC CLUMPS OF PIGMENT IN HAIR SHAFTS, AND AN ACCUMULATION OF
 CC MELANOSOMES IN MELANOCYTES. MOST PATIENTS ALSO DEVELOP AN
 CC UNCONTROLLED T LYMPHOCYTE AND MACROPHAGE ACTIVATION SYNDROME,
 CC KNOWN AS HEMOPHAGOCYTIC SYNDROME, LEADING TO DEATH IN THE ABSENCE
 CC OF BONE MARROW TRANSPLANTATION.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 6 IQ DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 DILUTE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U90942; AAD00702.1; -;
 CC EMBL: Y07759; CAA69035.1; -;
 CC EMBL: Y07759; CAA69036.1; -;
 CC EMBL: 222957; CAA80533.1; -;
 CC EMBL: S74799; AAB3211.1; -;
 CC EMBL: AF055459; AAC14188.1; -;
 CC HSSP: P08799; 1MND.
 CC MIM: 160777; -;
 CC MIM: 214450; -;
 CC InterPro: IPR002710; DIL.
 CC InterPro: IPR000048; IQ.
 CC InterPro: IPR001609; myosin_head.
 CC Pfam: PF01843; DIL; 1.
 CC Pfam: PF00612; IQ; 6.
 CC Pfam: PF00063; myosin_head; 1.
 CC PRINTS: PR00193; MYOSINHEAVY.
 CC PRODOM: PD000355; myosin_head; 1.
 CC PRODOM: PD003376; DIL; 1.
 CC SMART: SM00015; IQ; 6.
 CC SMART: SM00242; MYSC; 1.
 CC PROSITE: PSS0096; IQ; 6.
 CC MYosin: Repeat: ATP-binding; Calmodulin-binding; Actin-binding;
 CC Coiled coil: Phosphorylation; Alternative splicing; Polymorphism.
 CC Colled coil: 1 765 MYOSIN HEAD-LIKE.
 CC FT DOMAIN 1 766 788 IQ 1.
 CC FT DOMAIN 789 818 IQ 2.
 CC FT DOMAIN 819 836 IQ 3.
 CC FT DOMAIN 837 861 IQ 4.
 CC FT DOMAIN 862 883 IQ 5.
 CC FT DOMAIN 885 914 IQ 6.
 CC FT 914 1237 COILED COIL (POTENTIAL).
 CC FT DOMAIN 1338 1445 COILED COIL (POTENTIAL).
 CC FT

FT DOMAIN 1687 1792 DILUTE.
 FT NP_BIND 163 170 ATP (POTENTIAL).
 FT DOMAIN 643 665 ACTIN-BINDING (POTENTIAL).
 FT MOD_RES 1760 1760 PHOSPHORYLATION (POTENTIAL).
 FT VARSPPLIC 1321 1347 MISSING (IN ISOFORM 2).
 FT VARSPPLIC 1413 1413 MISSING (IN ISOFORM 3).
 FT VARSPPLIC 1413 1413 R -> C.
 FT VARIANT 1246 1246 R -> C.
 FT CONFLICT 668 668 F -> L (IN REF. 1).
 FT CONFLICT 833 833 MISSING (IN REF. 4).
 FT CONFLICT 863 863 E -> G (IN REF. 1).
 FT CONFLICT 922 922 H -> R (IN REF. 1).
 FT CONFLICT 1061 1061 V -> L (IN REF. 5).
 FT CONFLICT 1089 1089 E -> Q (IN REF. 4).
 FT CONFLICT 1177 1177 D -> E (IN REF. 5).
 FT CONFLICT 1465 1477 NIPKEDPQGM -> SYLCACGAVTVR (IN REF. 4).
 FT CONFLICT 1471 1471 K -> N (IN REF. 5).
 FT CONFLICT 1484 1484 E -> D (IN REF. 5).
 FT SEQUENCE 1855 AA; 215419 MW; 1C55AD57285FA9EC CRC64;
 SQ

Query Match 11.8%; Score 66; DB 1; Length 1855;
 Best Local Similarity 37.0%; Pred. No. 72; Mismatches 24; Indels 0; Gaps 0;
 Matches 17; Conservative

QY 19 VNGCVSTGVNARKEDNOQTIEQTIKTKNQKQETSRFGSADISF 64
 DB 183 VGSASANEVNEKVLTSNPMESIGNAKTRNDNSRGRKYEIGF 228

RESULT 12
 TRPD_SULSO STANDARD; PRT; 345 AA.
 ID TRPD_SULSO
 AC P50384;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Anthranilate phosphoribosyltransferase (EC 2.4.2.18).
 GN TRPD OR S500890.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 5833 / WT-4;
 RC Tutino M.L., Cubellis M., Sanna G., Marino G.;
 RT "The tryptophan operon in Sulfolobus solfataricus.";
 RT Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=ATCC 35092 / DSM 1617 / P2;
 RC MEDLINE=21195237; PubMed=11298741;
 RX She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Awayez M.J., Chan-Welher C.C.-Y., Clausen I.G., Cutlis B.A.,
 RA De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doellittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
 RA "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 RL [3]
 RN CHARACTERIZATION, AND CRYSTALLIZATION.
 RP MEDLINE=21195237; PubMed=11298741;
 RX Ivens A., Mayans O., Szadkowski H., Wilmanns M., Kirschner K.;
 RA "Purification, characterization and crystallization of thermostable
 RT anthranilate phosphoribosyltransferase from Sulfolobus solfataricus.";
 RL Eur. J. Biochem. 268:2246-2252(2001).
 CC -1- CATALYTIC ACTIVITY: Anthranilate + phosphoribosylpyrophosphate -
 CC N-5'-phosphoribosyl-anthranilate + diphosphate.
 CC -1- PATHWAY: SECOND STEP IN BIOSYNTHESIS OF TRYPTOPHAN.

```

CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE ANTHRANILATE PHOSPHORIBOSYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z50014; CA90309.1; -.
CC EMBL: AE006710; AA41173.1; -.
CC InterPro: IPR000312; Glycos_transf_3.
CC Pfam: PF00591; Glycos_transf_3; 1.
CC Pfam: PF02885; Glycos_transf_3; 1.
CC ProDom: PD001864; Glycos_transf_3; 1.
CC TrpCophan biosynthesis; Transferase; Glycosyltransferase;
CC Complete proteome.
CC SEQUENCE 345 AA; 37573 MW; D51927F4B7AAFA90 CRC64;

Query Match 11.7%; Score 65.5; DB 1; Length 345;
Best Local Similarity 21.2%; Pred. No. 13;
Matches 28; Conservative 21; Mismatches 42; Indels 41; Gaps 5;

OY 8 LTISALLTALLVVG-----CVSTGVNANKEQNOQTTEQ 40
DB 36 ILVSAIVLVAALRMKESKNEIVGFARAMELAKIDVPAINDTACTGDDGLCTGVVSTASA 95
OY 41 TIIR-----GKTNKQETISFRGSADISFMIVIKFGHTAILAPNMOETLSL-ISPLW 94
DB 96 ILTSLVPAVKHGRNAVSGKSGSD-----VLPAIGYINIIYPPRAKELVKNKINPFLF 149
OY 95 VKPYR--KNLS 104
DB 150 AQYHYPMKNVA 161

RESULT 13
SECT_CYAPA STANDARD; PRT; 492 AA.
ID SECT_CYAPA
AC P25014;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Preprotein translocase secy subunit.
GN SECY.
OS Cyanophora paradoxa.
OC Cyanelle.
OC Eukaryota: Glaucocystophyceae: Cyanophoraceae: Cyanophora.
OX NCBI_TaxID=2762;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91117189; PubMed-2126059;
RA Michalowski C.B., Pfanzagl B., Loeffelhardt W., Bohnert H.J.;
RT "The cyanelle S10 spc ribosomal protein gene operon from Cyanophora
RT paradoxa.";
RL Mol. Genet. 224:222-231(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-LB555 / PRINGSHEIM;
RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";
RL Plant Mol. Biol. Rep. 13:327-332(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-LB555 / PRINGSHEIM;
RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
RA Farley J.Y., Schluchter W.M., Chung J., Newman-Spallart C.,
RA Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;

```

```

RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
RT the genetic complexity of a primitive plastid.";
RL (in) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
RL Schweinler W. (eds.);
RL Eukaryotism and Symbiosis, pp.40-48; Springer-Verlag, Heidelberg
RL (1997).
CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT. PROBABLY INTERACTS WITH
CC OTHER PROTEINS TO ALLOW THE TRANSLLOCATION OF PROTEINS ACROSS THE
CC CYANELLE ENDOPLASMIC RETICULUM (CER) MEMBRANS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. CYANELLE.
CC -1- SIMILARITY: BELONGS TO THE SECY/SEC61-ALPHA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M30487; AAA63630.1; -.
CC EMBL: U30821; AAA81218.1; -.
CC Mendel: 7931; CYAPA; secy; 1.
CC InterPro: IPR002208; Secy.
CC Pfam: PF00344; secy; 1.
CC PRINTS: PR00303; SECYTRNLCASE.
CC PROSITE: PS00755; SECY_1; 1.
CC PROSITE: PS00756; SECY_2; 1.
CC Protein transport; Transmembrane; Cyanelle; Translocation.
CC SEQUENCE 492 AA; 55926 MW; 8736219112EBFCOD CRC64;

Query Match 11.6%; Score 64.5; DB 1; Length 492;
Best Local Similarity 41.4%; Pred. No. 24;
Matches 12; Conservative 7; Mismatches 9; Indels 1; Gaps 1;

OY 82 WOETLSLIISFLWVKPYRKNLSFYLTAK 110
DB 161 WAFIQSTIVISWIRPY-ALNMDFLGLK 188

RESULT 14
Y406_MYCPN STANDARD; PRT; 157 AA.
ID Y406_MYCPN
AC O50325;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG406 homolog (C12_orf157L).
GN MPN605 OR MP237.
GN Mycoplasma pneumoniae.
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae: Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 29342 / M129;
RX MEDLINE-96177562; PubMed-8604303;
RA Hilbert H., Himmelreich R., Plagens H., Herrmann R.;
RT "Sequence analysis of 56 kb from the genome of the bacterium
RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and a
RT cluster of ribosomal protein genes.";
RL Nucleic Acids Res. 24:628-639(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 29342 / M129;
RX MEDLINE-97105885; PubMed-8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC

Search completed: July 30, 2002, 15:42:00
Job time: 184 sec

THIS PAGE BLANK (US...)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:38:56 ; Search time 24.69 seconds
(without alignments)
777.741 Million cell updates/sec

Title: US-10-048-197-2
Perfect score: 558
Sequence: 1 MNNHRIRLTSLTALLVT.....FLWKPYPKNSLYLTAKA 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	78	14.0	299	16	Q97N22	Q97N22 streptococ
2	75	13.4	1195	13	Q9PUM2	Q9PUM2 xenopus lae
3	74	13.3	458	2	Q9L5C3	Q9L5C3 lactobacilli
4	73	13.1	112	16	Q9JW76	Q9JW76 neisseria m
5	72.5	13.0	138	16	P73777	P73777 synchocyst
6	71	12.7	568	10	Q9LXES	Q9LXES arabidopsis
7	69.5	12.5	328	15	Q9IU91	Q9IU91 human immun
8	69	12.4	168	15	Q9O1A9	Q9O1A9 human immun
9	69	12.4	530	5	Q01537	Q01537 caenorhabdi
10	68.5	12.3	169	2	Q9KZ55	Q9KZ55 streptomyce
11	68.5	12.3	380	5	Q00947	Q00947 tetrahymena
12	68.5	12.3	391	16	Q84574	Q84574 chlamydia t
13	68	12.2	182	6	Q77772	Q77772 oryctolagus
14	68	12.2	325	10	Q94B61	Q94B61 arabidopsis
15	68	12.2	360	15	P87977	P87977 human immun
16	68	12.2	430	16	Q9CE58	Q9CE58 lactococcus

17	68	12.2	1435	11	Q9J1Z2	Q9J1Z2 mus musculu
18	68	12.2	1444	11	Q9JLP0	Q9JLP0 mus musculu
19	68	12.2	1444	11	Q9J1Z1	Q9J1Z1 mus musculu
20	68	12.2	1445	11	Q9JW81	Q9JW81 mus musculu
21	68	12.2	1454	11	Q9JW82	Q9JW82 mus musculu
22	68	12.2	1455	11	Q9JW80	Q9JW80 mus musculu
23	67.5	12.1	210	16	Q9KD84	Q9KD84 bacillus ha
24	67	12.0	366	15	Q71102	Q71102 human immun
25	67	12.0	455	5	Q19553	Q19553 caenorhabdi
26	67	12.0	488	17	Q57576	Q57576 methanococ
27	67	12.0	799	16	Q9CH57	Q9CH57 lactococcus
28	67	12.0	819	16	Q31582	Q31582 bacillus su
29	67	12.0	873	15	Q90062	Q90062 human immun
30	66.5	11.9	310	16	Q97D49	Q97D49 clostridium
31	66.5	11.9	364	10	Q39162	Q39162 arabidopsis
32	66.5	11.9	364	10	Q9LX25	Q9LX25 arabidopsis
33	66.5	11.9	429	16	Q9ZM38	Q9ZM38 helicobacte
34	66.5	11.9	754	13	Q13131	Q13131 oncorhynch
35	66.5	11.9	754	13	Q13132	Q13132 oncorhynch
36	66.5	11.9	766	9	Q80245	Q80245 mycoplasma
37	66	11.8	396	10	Q64545	Q64545 arabidopsis
38	66	11.8	421	4	Q9BFX7	Q9BFX7 homo sapien
39	66	11.8	421	4	Q9H779	Q9H779 homo sapien
40	66	11.8	458	16	Q9ZDC5	Q9ZDC5 rickettsia
41	66	11.8	726	3	Q9HDP5	Q9HDP5 podospira a
42	65.5	11.7	103	2	Q46690	Q46690 escherichia
43	65.5	11.7	161	4	Q9Y3V4	Q9Y3V4 homo sapien
44	65.5	11.7	182	17	Q29802	Q29802 archaeoglob
45	65.5	11.7	199	10	Q9S0U2	Q9S0U2 arabidopsis

ALIGNMENTS

RESULT	ID	Q97N22	PRELIMINARY:	PRT:	299 AA.
AC	Q97N22	Q97N22			
DT	01-OCT-2001	(TREMBLREL, 18, Created)			
DT	01-OCT-2001	(TREMBLREL, 18, Last sequence update)			
DE	01-DEC-2001	(TREMBLREL, 19, Last annotation update)			
DE	CATION EFFLUX SYSTEM PROTEIN.				
GN	SP1857.				
OS	Streptococcus pneumoniae.				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_TaxID=1313;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=TIGR4;				
RC	MEDLINE=21357209; PubMed=11463916;				
RA	Tetzelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,				
RA	Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,				
RA	Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,				
RA	Umayam L.A., White O., Salzberg S.L., Lewis M.R., Ralston D.,				
RA	Holtzapple E., Khouri H., Wolf A.M., Uterback T.R., Hansen C.L.,				
RA	McDonald L.A., Feldblyum T.V., Anguoli S., Dickinson T., Hickey E.K.,				
RA	Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,				
RA	Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.,				
RT	"Complete genome sequence of a virulent isolate of Streptococcus				
RT	pneumoniae."				
RL	Science 293:498-506(2001).				
DR	EMBL: AE007477; AAK75929.1; .				
DR	TIGR: SP1857.				
DR	InterPro: IPR002524; Cation_efflux.				
DR	Pfam: PF01545; Cation_efflux; 1.				
KW	Complete proteome.				
SO	SEQUENCE 299 AA; 33528 MW; 62C23A92323E23DB CRC64;				

Query Match 14.0% Score 78; DB 16; Length 299;
Best Local Similarity 28.3% Pred. No. 2.6;
Matches 36; Conservative 18; Mismatches 39; Indels 34; Gaps 6;

RL Mature 404:502-506(2000).
 DR EMBL: AL162753; CAB83802.1; -
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 112 AA; 11515 MW; E5A9D201BF1D168A CRC64;

Query Match 13.1%; Score 73; DB 16; Length 112;
 Best Local Similarity 31.2%; Pred. No. 3;
 Matches 20; Conservative 9; Mismatches 35; Indels 0; Gaps 0;

OY 7 RLTSALTALLVTCGVTGVAMKEQNTIEOTIIGKTNKOEISRFSGSADISFMI 66
 Db 4 RLVSAAFEVALALAGCSINNVTVSDKIQERAAFGALGVSPNAVKISNRSGIRINETA 63

OY 67 VVIR 70
 Db 64 TVGR 67

RESULT 5
 P73777

ID P73777 PRELIMINARY; PRT; 138 AA.

AC P73777;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE HYPOTHEICAL 14.6 KDA PROTEIN.
 GN SL1158.
 OS Synecocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kocani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hiroseawa M., Sugitara M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL: D90909; BAA17829.1; -
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 138 AA; 14585 MW; 7753B6310931D13C CRC64;

Query Match 13.0%; Score 72.5; DB 16; Length 138;
 Best Local Similarity 32.0%; Pred. No. 4.2;
 Matches 24; Conservative 11; Mismatches 23; Indels 17; Gaps 3;

OY 3 NHHRLTISALLTALVTCGVTGVAMKEQNTIEOTIIGKTNKOEISRFSGSADISFMI 45
 Db 12 NSMRRLKVVSLTPPLIAGMIGLDAVLSFLPQTOSTAIAAPPASQOEIETLIASKG 71

OY 46 KTNKOEISRFSGSAD 60
 Db 72 KTGSDDLRRFFYGD 86

RESULT 6
 O9LXF5

ID O9LXF5 PRELIMINARY; PRT; 568 AA.

AC O9LXF5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHEICAL 62.0 KDA PROTEIN.
 GN F8M21_160.
 OS Arabidopsis thaliana (mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
 RA Rudd S., Lemckey K., Mayer K.F.X.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL353933; CAB89337.1; -
 DR InterPro: IPR004087; KH.
 DR InterPro: IPR004088; KH-domain; 4.
 DR Pfam: PF00013; KH-domain; 4.
 DR SMART: SM00322; KH; 4.
 DR PROSITE: PS50084; KH-type_1; 4.
 KW Hypothetical protein.
 SQ SEQUENCE 568 AA; 62034 MW; 3431FB92B4E37205 CRC64;

Query Match 12.7%; Score 71; DB 10; Length 568;
 Best Local Similarity 27.1%; Pred. No. 30;
 Matches 23; Conservative 14; Mismatches 34; Indels 14; Gaps 3;

OY 15 TALLV---TCGV-STGVAMKEQNTIEOTIIGKTNKOEISRFSGSADISFMIYVI 69
 Db 372 TRLVPSRRICIGLGGATITTEMRMTKANIRILGKELPKVASDDDEMOVNMVLL 431

OY 70 KFGHTAILAPNRMCIELSLISFLW 94
 Db 432 KFS-----LQFLSLRLRFETW 447

RESULT 7
 O9IU91

ID O9IU91 PRELIMINARY; PRT; 328 AA.

AC O9IU91;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE EXTERIOR MEMBRANE PROTEIN GP120 (FRAGMENT).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=96NG-MACSW031;
 RX MEDLINE=20179113; PubMed=10716369;
 RA Peeters M., Esu-Williams E., Vergne L., Montavon C.,
 RA Mulanga-Kabeya C., Harry T., Iblironke A., Lesage D., Patrel D.,
 RA Delaporte E.;
 RT "Predominance of subtype A and G HIV type 1 in Nigeria, with
 RT geographical differences in their distribution.";
 RL AIDS Res. Hum. Retroviruses 16:315-325(2000).
 DR EMBL: AJ389751; CAB77485.1; -
 DR InterPro: IPR007777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR AIDS; Coat protein; Glycoprotein.
 KW AIDS; Coat protein; Glycoprotein.
 FT NON_TER 1
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 328 AA; 36504 MW; DE92C03A75CF898E CRC64;

Query Match 12.5%; Score 69.5; DB 15; Length 328;
 Best Local Similarity 21.2%; Pred. No. 23;
 Matches 25; Conservative 29; Mismatches 43; Indels 21; Gaps 6;

OY 4 HHRLTISALLTALVTCGVTGVAMKEQNTIEOTIIGKTNKOE 51
 Db 120 HGKRPVVS---TOLLNGLAEEDIVRTENTYDNAKIIVOLNETIENCTRPNNMTRK 176

OY 52 ISSRFGSADSI---SFMIVIKFGHTAILAPNRKOEILSLI-ISPLWYKPYRKNLSF 105
 DB 177 -STRFPGQAFYATGDIENIRQHCNV-SRTKWMEMIOKRVOLRKIESFNKNISF 232

RESULT 8
 ID Q901A9 PRELIMINARY; PRT: 168 AA.
 AC Q901A9;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE GP120 (FRAGMENT).
 OS Human immunodeficiency virus type 1.
 ON ENV.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_Taxid=11676;
 RN [1]
 RA SEQUENCE FROM N.A.
 RA Esteves A., Parreira R., Piedade J., Veneno T., Franco M., Sousa G.,
 RA Canas-Ferreira W.F.;
 RT "Molecular epidemiology of HIV-1 in Portugal: high prevalence of non-B
 RT subtypes in Lisbon."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ296262; CAC82697.1;
 FT NON_TER 1
 FT 168
 FT SEQUENCE 168 AA; 18373 MW; AEF105965BCA53B CRC64;

Query Match 12.4%; Score 69; DB 15; Length 168;
 Best Local Similarity 27.1%; Pred. No. 12;
 Matches 32; Conservative 20; Mismatches 42; Indels 24; Gaps 6;

OY 4 HIRLTISALLTALVTGCVSTGNVAMKEONQOTIEOTII--KGKTKKOEISSRFGSAD 60
 DB 42 HGKRPVVS---TQLLNSGLADGNTIRSENTNNAKIIIVOLKEAVNITCIRPSNMTRK 98
 OY 61 SISF-----MIVIKFGHTAILAPNR--WOELSLIISPLWYKPYRKNLSF 105
 DB 99 SISFPGQAFYATGDIENIRQHCNV--SRTKWMEMIOKRVOLRKIESFNKNISF 151

RESULT 9
 ID Q001537 PRELIMINARY; PRT: 530 AA.
 AC Q001537;
 DT 01-JUL-1997 (TREMblrel. 04, Created)
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE HYPOTHETICAL 58.8 KDA PROTEIN.
 GN F21F8.11.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_Taxid=6239;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-BRISTOL N2;
 RC MEDLINE=99069613; PubMed=9851916;
 RX None;
 RA "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium."
 RT Science 282:2012-2018(1998).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-BRISTOL N2;
 RC Wilson R., Favello A., Le T.T.;
 RA "The sequence of C. elegans cosmid F21F8."
 RT Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN-BRISTOL N2;

RA Waterston R.;
 RT "Direct Submission."
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: U97000; AAC47997.1;
 DR InterPro: IPR003662; sub.transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR Hypothetical protein; Transmembrane.
 KW SEQUENCE 530 AA; 58783 MW; FBDOC3DA769B78A6 CRC64;

Query Match 12.4%; Score 69; DB 5; Length 530;
 Best Local Similarity 21.7%; Pred. No. 45;
 Matches 30; Conservative 25; Mismatches 47; Indels 36; Gaps 5;

OY 5 HIRLTISALLTALVTGCVSTGNVAMKEONQOTIEOTIIKGKTKKOEISSRFGSAD---- 60
 DB 44 HIGLSLCMCNSTAV-ALMTNNTNATVEGTESTIMSTLENKTSF---SQELGGDGGES 99
 OY 61 -----SISFMIIVIK-----FGHTAILAPNRW--QELSLIISF 92
 DB 100 CTKLEKVIKDYGGFTMSVSMOGIYSAFLGGRIFEPAGVLVDRFSARHILSVAILM 159
 OY 93 LWKPYRPNLSFYLAK 110
 DB 160 LTIASLMPVLSTYIGEK 177

RESULT 10
 ID Q9KZ25 PRELIMINARY; PRT: 169 AA.
 AC Q9KZ25;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE PUTATIVE INTEGRAL MEMBRANE PROTEIN.
 GN SCD6.09C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_Taxid=1902;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-A3(2);
 RC Oliver K., Harris D.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RC Cerdano A.M., Parkhill J., Barrell B.G., Randsdram M.A.;
 RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN-A3(2);
 RC MEDLINE=97000351; PubMed=8843436;
 RX Redenbach M., Kleser H.M., Denapalite D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL353815; CAB88458.1;
 DR InterPro: IPR000566; Lipocalin_cyFABP.
 DR PROSITE: PS00213; LIPOCALIN; UNKNOWN_1.
 SEQUENCE 169 AA; 18814 MW; B7685780ECA02D77 CRC64;

Query Match 12.3%; Score 68.5; DB 2; Length 169;
 Best Local Similarity 24.0%; Pred. No. 14;
 Matches 24; Conservative 20; Mismatches 31; Indels 25; Gaps 5;

OY 8 LTISALLTALIV--TGCSTGNVAMKEONQOTIEOTIIKGKTKKOE-----ISSRFGSA 59

Db 10 LLAATVLAVALALVALGNITDFDNNQYVHNVLAAMDTEFKDDMLMRATSK-GLQ 68
 QY 60 DSISFIVIKFGHTAILAPNRMOELSLIIT---SPLMWK 96
 Db 69 DTAVVAIV-----WETAALVILYIMATWLMIR 95

RESULT 11
 ID 000947 PRELIMINARY; PRT: 380 AA.
 AC 000947;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE TREMBSKELETAL (FRAGMENT).
 OS Tetrahymena vorax.
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
 OC Tetrahymenina; Tetrahymena.
 ON NCBI_TaxID=5919;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CL;
 RA Green M.;
 RT "Tetraskeleal: A unique constitutively expressed structural protein
 of Tetrahymena vorax.";
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF003091; AAB61407.1; -
 DR InterPro: IPR003015; HLH_Myc.
 DR POSITIVE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 380 AA; 41058 MW; 4F58C8DF5444E9A3 CRC64;

Query Match 12.38; Score 68.5; DB 5; Length 380;
 Best Local Similarity 41.58; Pred. No. 35;
 Matches 17; Conservative 7; Mismatches 16; Indels 1; Gaps 1;
 QY 9 TISALLTALVTGCVSTGVNANKNOQTIETQITIKGKTNK 49
 Db 282 TINTMTAMATTITIDMTNGVTKEEN-PTIVRTTEANTNK 321

RESULT 12
 ID 084574 PRELIMINARY; PRT: 391 AA.
 AC 084574;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE GEN. SECRETION PROTEIN F.
 GN GSPF OR CT570.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 ON NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-D/UM-3/CX;
 RX MEDLINE=99000809; PubMed=9784136.
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA Mitchell W.P., Olinger L., Tatusev R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans:
 Chlamydia trachomatis.";
 RL Science 282:754-759(1998).
 DR EMBL: AE001327; AAC68172.1; -
 DR InterPro: IPR001992; Bact_scat_systII.
 DR Pfam: PF00482; GSP1_F.1.
 KW Complete proteome.
 SQ SEQUENCE 391 AA; 43557 MW; 7785601AAB540650 CRC64;

Query Match 12.38; Score 68.5; DB 16; Length 391;
 Best Local Similarity 21.78; Pred. No. 36;

Matches 20; Conservative 22; Mismatches 39; Indels 11; Gaps 3;
 QY 14 LTALLVTCVSTGVNANKNOQTIETQITIKGKTNKQISSRFSADSISPMI----- 66
 Db 275 LIEALTLCCEAVSODPFRKEELQEVY-QAVVNGSLSRHSHRTWPKVLIGVALGEBG 333
 QY 67 -VVIKFGHTAILAPNRMOELSLIISFLWKP 97
 Db 334 DLAVFAHVAQIVNEDIQRLTWTA--WCOP 363

RESULT 13
 ID 077772 PRELIMINARY; PRT: 182 AA.
 AC 077772;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUTATIVE PLASMINOGEN ACTIVATOR INHIBITOR-1 (FRAGMENT).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 ON NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA Yin J., Idell S.;
 RT "Cloning and Expression of Rabbit PAI-1,"
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR EMBL: AF074325; AAC62632.1; -
 DR HSSP: P05121; 1A7C.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR SMART: SM00093; SERPIN.1.
 KW Serpin.
 FT NON_TER 1
 SQ SEQUENCE 182 AA; 20920 MW; B2B6AEB6D92E217F CRC64;

Query Match 12.28; Score 68; DB 6; Length 182;
 Best Local Similarity 25.08; Pred. No. 17;
 Matches 26; Conservative 13; Mismatches 41; Indels 24; Gaps 3;

QY 7 RLTSALLTALVTGCVSTGVNANKNOQTIETQITIKGKTNKQISSRFSADSISPMI 66
 Db 82 QLTLLLVNALYFNGQWKT-----PFSKSGTHHRYFHKSDSTISVPMYA 126
 QY 67 VVIKFGHTAILAPNRMOELSLIISFLWKPYPKKNLSFYLA 109
 Db 127 QTNKFNTTEFLTPDGHYDIEL-----PYHGETLSMFLAA 162

RESULT 14
 ID 094B61 PRELIMINARY; PRT: 325 AA.
 AC 094B61;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE UNKNOWN PROTEIN.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 ON NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Bann J., Garinici P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamuya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:39:21 ; Search time 30.09 Seconds

(without alignments)
409.744 Million cell updates/sec

Title: US-10-048-197-2

Perfect score: 111
Sequence: 1 MLNHHIRLTISALLTALLVT.....FLWVKYRPNKLSFYLTAKA 111Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 15

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : A.Geneseq_032802.*

```
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*
```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	100.0	111	22 AAB60640	Moraxella catarrhalis

ALIGNMENTS

RESULT 1
AAB60640
ID AAB60640 standard: Protein: 111 AA.
XX
AC AAB60640;

```
XX 02-MAY-2001 (first entry)
DT Moraxella catarrhalis strain ATCC43617 BASB122 protein.
XX
XX
DE BASB122 protein: strain ATCC43617; antigen; antibody; vaccine;
XX genetic immunisation; infection; upper respiratory tract; otitis media;
XX hearing loss; deafness; pneumonia; sinusitis; nosocomial infection;
XX invasive disease; antibacterial; auditory.
XX
OS Moraxella catarrhalis.
XX
XX WO200109337-A2.
XX
XX 08-FEB-2001.
XX
XX 31-JUL-2000; 2000WO-EP07365.
XX
XX 30-JUL-1999; 99GB-0018034.
XX
XX 30-JUL-1999; 99GB-0018036.
XX
XX (SMIT) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Thonard J;
XX
XX WPI: 2001-159874/16.
XX
XX N-PSDB: AAF59778.
XX
XX New BASB122 and BASB124 polypeptides and polynucleotides from Moraxella
XX catarrhalis strain ATCC 43617, useful as therapeutic agents or vaccines
XX against bacterial infections, e.g. otitis media or pneumonia
XX
XX Claim 4; Page 66; 75pp; English.
XX
XX The invention relates to the Moraxella catarrhalis strain ATCC43617
XX BASB122 and BASB124 proteins (AAB60640 and AAB60641, respectively)
XX and to DNA encoding them (AAF59778 and AAF59779, respectively). The
XX invention also relates to immunogenic fragments of the BASB122 and
XX BASB124 proteins, expression vectors and host cells comprising BASB122
XX or BASB124 nucleic acids, the recombinant production of BASB122 or
XX BASB124, vaccine compositions comprising the BASB122 or BASB124 proteins
XX or nucleic acids, an antibody against BASB122 or BASB124, therapeutic
XX compositions comprising an anti-BASB122 or anti-BASB124 antibody, and a
XX method of identifying a Moraxella catarrhalis infection via the
XX detection of BASB122 or BASB124 proteins or antibodies. The vaccine
XX compositions of the invention are useful as prophylactic or therapeutic
XX agents against Moraxella catarrhalis infections in mammals, particularly
XX humans. Moraxella catarrhalis is a Gram negative bacterium frequently
XX isolated from the human upper respiratory tract, which is responsible for
XX several pathological conditions. It is responsible for about 15% of
XX otitis media cases in children (which can lead to temporary or permanent
XX hearing loss). It also causes pneumonia in elderly people, and sinusitis,
XX nosocomial infections and, less frequently, invasive diseases. BASB122 or
XX BASB124 proteins or nucleotides may additionally be used in screening for
XX novel antibacterial compounds, and in the diagnosis and staging of
XX infections. The present sequence represents the Moraxella catarrhalis
XX strain ATCC43617 BASB122 protein.
XX
XX Sequence 111 AA:
SQ
```

Query Match 100.0%; Score 111; DB 22; Length 111;
Best local similarity 100.0%; Pred. No. 3.8e-111;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MLNHHIRLTISALLTALLVTCVSTGVNVMKEQNOOTIEQTIIRKGTNRQEISSRFGSAD 60
D 1 mlnhhirltisalltallvtcsvgtnvmkeqngqtleqtlkgytknqgsrfgsad 60
QY 61 SISEFMVIVIKFGHTALAPNRQELSLTISFLWVKYRPNKLSFYLTAKA 111
D 61 sisefmviviqhtalaprwnqelssltisflwvkrpknlsfylvltaka 111
```

Wed Jul 31 08:21:27 2002

Search completed: July 30, 2002, 15:43:08
Job time: 227 sec

us-10-048-197-2.oligo.rag

Page 2

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:39:51 ; Search time 13.05 Seconds
(without alignments)
207.758 Million cell updates/sec

Title: US-10-048-197-2
Perfect score: 111
Sequence: 1 MLNHHIRLTISALLTALVT.....FLWVKPYRPKNLSFYLTAKA 111

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

No matches found

Search completed: July 30, 2002, 15:43:27
Job time: 216 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:41:46 ; Search time 16.2 Seconds
(without alignments)
658.390 Million cell updates/sec

Title: US-10-048-197-2

Perfect score: 111

Sequence: 1 MLNHHIRLTISALLTALLVT.....FLWVKPYRPNLSFYLTAKA 111

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR_71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match length	ID	Description
---------------	----------------	--------------	----	-------------

No matches found

Search completed: July 30, 2002, 15:43:50
Job time: 124 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OK protein - protein search, using sw model

Run on: July 30, 2002, 15:43:31 ; Search time 11.9 Seconds
(without alignments)
361.166 Million cell updates/sec

Title: US-10-048-197-2

Perfect score: 111

Sequence: 1 MLNHHIRLTISALLTALVT.....FLWKKPYRPKNLSFYLTAKA 111

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 105224 segs, 38719550 residues

Word size : 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	ID	Description
No matches found			

Search completed: July 30, 2002, 15:47:02
Job time: 211 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:43:11 ; Search time 24.65 Seconds
(without alignments)
779.004 Million cell updates/sec

Title: US-10-048-197-2
Perfect score: 111
Sequence: 1 MLNHHIRLTISALLTLVLT.....FLWKPYRPNLSPLYTAKA 111

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
No matches found						

Search completed: July 30, 2002, 15:46:44
Job time: 213 sec

